Lancaster University Management School

Economics Working Paper Series

2020/013

The acceptability of lotteries in allocation problems: a choice-based approach

Elias Bouacida and Renaud Foucart

The Department of Economics
Lancaster University Management School
Lancaster LA1 4YX
UK

The acceptability of lotteries in allocation problems: a choice-based approach

Elias Bouacida[*] Renaud Foucart[†]

July 20, 2020

Abstract

We report the results of two experiments on the social acceptability of random devices in allocation mechanisms. A majority of subjects do not opt for a lottery if they can rationalize an alternative mechanism as non-random. It is, however, possible to design a payoff-equivalent mechanism to the lottery that is more acceptable. Our results shed light on the real-world reliance on obscure criteria in allocation problems where lotteries seem to be simpler and more efficient.

Keywords: lotteries, mechanism design

JEL-Code: D01, D78, D91

1 Introduction

Consider an allocation problem in which a subject can either receive a high or a low reward. This problem can be solved either with a lottery or with a set of criteria. If the two methods offer the same odds of getting the high reward, will the subject prefer one over the other? In this paper, we report the results of two experiments revealing a preference for the use of criteria. The preference is not explained by overconfidence but seems to originate from an aversion to the explicit use of a lottery. Reframing the lottery by allowing subjects to take a payoff-irrelevant decision makes it more acceptable.

Many of the solutions provided by economists to real-world allocation problems involve a role for randomization, usually as a tie-breaking rule. A leading example is the school allocation mechanisms, such as the deferred acceptance algorithm.[1] In practice, however, allocations problems solved using explicit lotteries are rare (we discuss some of them in Appendix A). For instance, in USA law, the Administrative Procedure Act (APA) forbids an agency to take any action that is "arbitrary, capricious, or an abuse of discretion" (Vermeule, 2015, p.475). It means that even in the absence of any information or rational reason to pick one decision over the other, a federal agency is compelled to provide a rationalization and can never toss a coin.[2] Anecdotal evidence from the implementation of school choice mechanisms in several countries

[*]Lancaster University Management School, e.bouacida@lancaster.ac.uk

[†]Lancaster University Management School, r.foucart@lancaster.ac.uk

[1]Debates on the use of randomization in school choice revolve on its impact on parents' strategies (Erdil and Ergin, 2008; Kesten and Ünver, 2015; Abdulkadiroglu *et al.*, 2019), but not on the acceptability of randomization in itself.

[2]What lawmakers find objectionable is not the fact that an agency has the same probability of reaching one decision or the other, but the failure to deliberate that the use of a randomization device represents (Leib and Galoob, 2016, p.1860). In a case of custody allocation, the supreme court of Virginia stated that "A judge's act of tossing a coin in a courtroom to decide a legal issue pending before the court suggests that courts do not decide cases on their merits but instead subject litigants to games of chance in serious matters without regard to the evidence or applicable law." (Supreme Court of Virginia, Record No. 071014).

also shows a reluctance of parents to accept lotteries and a preference for "objective" albeit often arbitrary and murky criteria.[3]

Another example is the award of research funding by peer review. Except for the very best and worst submissions, this procedure is largely unpredictable (Cole *et al.*, 1981; Graves *et al.*, 2011; Pier *et al.*, 2018). Given the high cost spent writing and evaluating proposals, it would be more cost-efficient to allocate funding randomly among applications deemed good enough, perhaps with a subset of outstanding proposals guaranteed funding (Greenberg, 1998; Brezis, 2007; Fang and Casadevall, 2016; Roumbanis, 2019; Avin, 2019). However, in practice, scientists and funding bodies have resisted this approach, insisting on providing a rationalization (Barnett, 2016).[4]

We ran two incentivized experiments on Amazon Mechanical Turk (AMT hereafter) to study whether there exists an individual revealed preference for the use of criteria over lotteries. Subjects were paid a show-up fee to participate and earned an additional reward based on the result of a game. A necessary step to be able to identify a potential preference was to identify a set of criteria for which we could provide exact probabilities of winning before subjects chose between it and a lottery. In both experiments, we asked subjects to give a sequence of 5 strategies of rock-paper-scissors (RPS) at the beginning. The choice of RPS strategies could not be modified further. It became a set of criteria for which the odds of winning against a given set of strategies could be computed. We then asked subjects to choose the process that determined if they won the reward. The first possibility was to use their chosen strategies against a player drawn from a data set of 2,500 real-world RPS players. They won the reward if they won more RPS games than their randomly drawn opponent. In case of ties, the first who won a round won the game.[5]

In experiment 1, the second possibility was to condition the reward on the toss of a fair coin by an algorithm (CT hereafter). We chose this device as the simplest and most familiar lottery, thereby more likely to be perceived as fair (Eliaz and Rubinstein, 2014). We find that, on average, 66% of the subjects chose the RPS over the CT. When we provided subjects their odds of winning RPS given their chosen strategies, some reacted to this information. On average, our regression analysis estimates that out of 100 subjects, 25 switch mechanism depending on whether or not their RPS strategies have a probability of winning higher than 1/2. Subjects knowing their odds of winning RPS were on average as likely to choose the coin toss that those who did not know it, showing that the preference for RPS is not explained by overconfidence. Moreover, subjects reacted similarly to the actual odds of winning and to their perceived probability of winning when deciding under uncertainty.

In experiment 2, on top of 5 RPS strategies, we asked subjects to provide 5 strategies of a repeated matching pennies (MP) game. The MP strategies would be played against a computer choosing head and tails with equal probability. We then offered subjects the same choice as in

[3]In France, the system of allocation of students to universities (APB) used lotteries to break ties when capacity was reached. It was criticized and is now replaced with another mechanism with a lot more criteria, reducing the likelihood of ties and ending the use of lotteries. Importantly, the weights between the criterion are not public and leave a lot of almost arbitrary leeway to the universities, which was not the case previously. In French-speaking Belgium, a 2009 attempt to randomly allocate students in oversubscribed high schools lasted only a year after being dubbed by some parents and the media the "lottery law" and replaced by a set of criteria still in use today. In the UK, when a lottery was implemented for school allocation in Brighton and Hove, the Conservative party (then in the opposition) pledged a country-wide ban of lotteries. On March 2008, before becoming Secretary of State for Education, Michael Gove justified this stance on the BBC as "Would you like the fate of your children to be decided by the spin of the roulette wheel or the roll of the dice? No, and neither would I. I think that one of the key problems with lotteries is that they reduce almost one of the most important decisions that any parent is going to make for their child, into a matter of chance." While the ban was never implemented, lotteries have not been extended either.

[4]The Health Research Council of New Zealand is a notable exception, running a pilot random allocation for its early career "Explorer" grants since 2013.

[5]And if there was still a tie, in the end, the player won if she had spent an even number of seconds choosing the RPS strategies.

experiment 1 but using MP instead of CT as an alternative to RPS. MP and CT are in practice equivalent. MP's odds are not influenced by a subject's choice of strategies. We find, however, that only 42% of our subjects chose the RPS in this setting, a proportion significantly different from experiment 1. Re-framing the lottery as the result of a choice, even when the expected payoff is the same, seems to make it much more socially acceptable. This result also suggests that the preference for RPS in experiment 1 is not based on that game being factually more "meritocratic", as the probability of winning the MP is independent of the chosen strategies. It also rules out the "competence hypothesis" of Heath and Tversky (1991) as an explanation. There is no reason subjects would feel relatively more competent in experiment 2 with MP than they are with CT in experiment 1.

The novelty of our approach is to provide an incentivized choice between a lottery and a set of criteria. Survey evidence shows that people are reluctant to rely on the use of a random device to determine the outcome of important hypothetical decisions involving other people (Keren and Teigen, 2010). Oberholzer-Gee *et al.* (1997) also report survey evidence that a market mechanism is the only procedure deemed less acceptable than a lottery for the allocation of a nuclear waste facility.

Our results relate to several strands in the literature. It is well-known that individuals have an "outcome bias" and a tendency to interpret success by merit and effort and ignore the role of luck (Frank, 2016; Brownback and Kuhn, 2019). Hence, subjects may be willing to interpret any device that is not explicitly a lottery as more meritocratic, even if it is completely random. Once a reward has been received for reasons perceived as more meritocratic, it is then valued more (Loewenstein and Issacharoff, 1994). This point is also related to the literature on "source uncertainty": Individuals treat uncertainty differently depending on the mechanism generating it (see, for instance, Heath and Tversky (1991); Fox and Tversky (1995)). Finally, there is an intrinsic value in holding decision rights (Bartling *et al.*, 2014; Bobadilla-Suarez *et al.*, 2017). It is consistent with our subjects interpreting the use of their RPS strategies as something they have "decided", and may explain why the MP in experiment 2 is more acceptable than CT in experiment 1.

Our result of an apparent aversion to being the *subject* of a lottery in an allocation problem contrasts with the way people behave when choosing between different options (see also Appendix A). Indeed, individuals often choose to rely on a randomization device to make their choice when indifferent or indecisive between two alternatives (Agranov and Ortoleva, 2017; Qiu and Ong, 2017; Bouacida, 2019; Cettolin and Riedl, 2019). This "preference for randomization" has been observed on subjects choosing strategies in allocation mechanisms such as school choice (Dwenger *et al.*, 2018), in which many subjects seem however reluctant for the mechanism itself to use lotteries.

In the next section, we detail the experimental protocol. We provide the results in Section 3 and discuss them in Section 4.

2 Experiment

We ran two experiments whose aim are to choose a mechanism for allocating a reward to yourself. The two mechanisms are a *lottery* with known probability and a *set of criteria.* The set of criteria is stylized to the choice of five actions of a repeated rock, paper, scissors game. The actions are then played against one player from a database of 2,500 we have compiled (the database is briefly described in Section 2.3). We draw the opponent with an equiprobable distribution. The subject should win more rounds than their opponent to win the RPS game. In case of a tie, the first round which is not a tie is used to determine the winner. If they have chosen the same five actions in the same order, then the subject wins if they spent an even number of seconds on the choice screen.

The lottery in experiment 1 is a coin toss. In experiment 2, the lottery consists of five choices of head and tail for a repeated matching penny game. The subject wins the MP whenever they match three or more head and tails drawn by the computer. The computer draws head or tail with equal probability. In both cases, the *ex ante* chances of winning the lottery are 50% and are independent of the chosen strategies.

The experiments were run on Amazon Mechanical Turk on April 29, 2020 (experiment 1) and May 13-14, 2020 (experiment 2), using oTree (Chen *et al.*, 2016). A total of 197 subjects finished experiment 1 and 89 experiment 2. The reward was $1.40 in experiment 1 and $0.40 in experiment 2. The show-up fee was $0.70 in experiment 1 and $0.40 in experiment 2. The median time spent in the experiments was 3 minutes. The median hourly wage was $25.71 in experiment 1 and $12.20 in experiment 2. The majority of our subjects declare themselves to be male (68.5% and 53.8%, respectively). They are mostly between 25 and 40 years old, employed and come from the United States. The two samples are similar to each other (see Appendix G for a more detailed discussion).

In both experiments, we elicit the subjects' beliefs have about their probability of winning in the RPS and the MP when these probabilities were not given. In both experiments, the belief elicitation procedure is incentivized. In experiment 1, it is done by asking the number of time they expect to win against our dataset (following Schlag and Tremewan, 2020). In experiment 2, we ask them the score they expect to reach in both RPS and MP (both scores are between -5 and 5 and are thus comparable). They are paid if their belief on their score is close enough from the real score.[6] The two belief elicitation methods are not equivalent, as the score in RPS does not exactly translate into a probability of winning.[7]

2.1 Timing

The experiments are divided into:

1. Description of the experiment;
2. Description of the RPS and MP games and choices of the five actions in each (only in experiment 2 for MP);
3. Choice of the allocation mechanism;
4. Incentivized belief elicitation, when appropriate, and asking subjects for their reason of choosing one mechanism over the other;
5. Demographic questions;
6. Incentivized ambiguity attitude elicitation (in experiment 1 only);
7. Feedback on their choices and payment.

2.2 Treatments

We have four different treatments. We call the first two treatments knowledge and ignorance. In the knowledge treatment, we tell the subjects their average probability of winning the RPS game before choosing the mechanism. That is, we tell them the percentage of RPS games of the database they win against. In the ignorance treatment, they receive no information about the probability of winning of their five RPS actions. We did not elicit their beliefs about the probability of winning the RPS in the knowledge treatment.

[6]The only correct belief in experiment 2 for MP is 0. In experiment 1, subjects beliefs were considered correct if they were at within 1/2500 of the correct belief. In experiment 2, if they were within 0.1 score of their real score.

[7]This is because of the tie-breaking rules.

We call the other two treatments chance and ability. The two treatments are two different framings of the RPS game. In the chance treatment, we tell the subjects that RPS is a game of chance, whereas, in the ability treatment, we tell them it is a game of ability and strategy. The exact framing is available in Appendix B. We are not deceiving the subjects. The two descriptions are accurate depictions of the game. Competitions of RPS exist and some players fare on average better, mainly by exploiting the biases of others (we describe the data sets in Appendix C). It is also a game of chance, as the Nash equilibrium strategy of this game is to play each action with probability one third. Playing Nash guarantees to win with 50% probability.

We have used a 2x2 factorial design for experiment 1 to assign the different treatments. The repartition of subjects in the different treatments is shown in Table 1. In experiment 2, subjects were primed to see the RPS as a game of chance and did not know their probability of winning. So experiment 2 was run using only ignorance and chance treatments.

Table 1: Number of subjects in each treatment in experiment 1.

	Ability	Chance
Knowledge	50	51
Ignorance	55	41

2.3 RPS Data

We have collected the 2,500 strategies of the opponent in the RPS game from actual games played on the website https://roshambo.me.[8] An overview of the data set is available in Table 8 in Appendix C, as well as a comparison with the observed choices in our two experiments. One of the well-known biases in RPS is that more than a third of players choose Rock as the first action, and less than a third choose scissors, which mean that playing Paper is the best first move. The average probability of winning against our dataset is always between 44.80% and 55.00%. In that sense, the set of criteria is closed from the lottery *ex ante*. The subjects do not know that. The only information about their chances of winning in the RPS game – in the Knowledge treatment – is the probability of winning against our dataset. Moreover, no strategy gives an exact probability of winning of 50%. As a consequence, a player valuing expected monetary payoff only should never be indifferent in the knowledge treatment.

3 Results

Most subjects choose the RPS in experiment 1 (66% overall treatments), whereas most of them choose the MP in experiment 2 (58%). The proportion of subjects choosing the RPS is significantly higher in experiment 1.[9] The proportion of subjects choosing the RPS is significantly different from 50% in experiment 1 but not in experiment 2.[10] In the next subsections, we explore possible explanations for these differences.

3.1 Effect of the Treatments

Table 2 shows that the different treatments do not seem to influence the choice of the allocation mechanism. What this similarity implies is that the observed preference for RPS cannot be

[8]We thank Lasse Hassing for giving us access to their data.

[9]The p-value of the Fisher exact test is < 0.001 between the two experiments. When restricting the sample of experiment 1 to the same treatments as experiment 2, the p-value is 0.023.

[10]The p-values of the one-sample two-sided t-test are < 0.001 and 0.14, respectively.

explained by overconfidence. Overall, subjects in the Knowledge treatment are equally likely to choose it than those in the Ignorance treatment. There is, however, a role for ambiguity aversion. We find that – as expected – ambiguity averse subjects are less likely to choose RPS than the others (see Appendix F for more detailed results).

Table 2: Proportion of subjects choosing each mechanism, depending on the treatment.

	Ability	Chance	Overall
Knowledge	64%	71%	67%
Ignorance	66%	65%	66%
Overall	65%	68%	66%

Note: the proportions are not significantly different between the different treatment cells (using a Fisher exact test).

Using this level of aggregation to understand choices is not sufficient. In expected monetary payoffs, the lottery is more valuable if the probability of winning the RPS is below 50%, and less valuable otherwise. When we introduce the 50% probability of winning the RPS threshold, it matters in the Knowledge treatment. In the Ignorance treatment, however, what should influence the subjects' choice is their belief that they are above or below 50%. Table 3 shows that it is indeed the case. When subjects know their probability of winning the 50% threshold makes a significant difference. It is also true when they believe it is the case. The real probability, however, does not matter when they do not know it. We illustrate this phenomenon graphically in Appendix D.

Table 3: Effect of knowing or believing that the chances of winning the RPS are above or below 50% in experiment 1.

	$q < 0.5$	$q > 0.5$	P-value[1]
Knowledge	58%	82%	0.020
Ignorance & Really better[2]	63%	70%	0.59
Ignorance & Believed better[3]	55%	80%	0.014

[1] P-value of the Fisher exact test;
[2] Really better means that the real probability of winning the RPS is above 50%;
[3] Believed better means that subjects believe their probability of winning the RPS is above 50%.

To summarize, a first variable that seems to matter to understand the aggregate choices in experiment 1 is whether they believe or know that their probability of winning the RPS is higher or lower than 50%. However, it matters only for a fraction of subjects, as 58% of them choose the RPS even when they know they have a lower probability of winning it, and 18% of them choose the CT when they have a higher probability of winning the RPS.

3.1.1 Regression Analysis

Table 4 report the result of a regression analysis that confirms the previous results.[11]The variable $RPS \succeq CT$ is a dummy that takes value 1 when subjects know or believe that their chance of

[11]We provide additional specifications in Appendix E. Table 9 justifies that we do not look at all interaction terms.

winning in the RPS is above 50%. Out of 100 subjects, an average of 25 choose RPS instead of CT when their probability of winning the RPS is above the threshold of 1/2. While it clearly shows that the probability of winning matters, it also implies that a majority of the subjects are not deciding only on their expected monetary payoff. As a rough breakdown, the estimation tells us that around 55% of the subjects would choose RPS even if it gives odds lower than 1/2, around 25% pick the device that gives them the highest probability of winning, and 20% choose CT even if the RPS has better odds. Note that the estimation only looks at the exact 50% threshold. Table 10 in Appendix E looks at the influence of the distance with the 50% threshold, instead of a dummy. The explanatory power is similar, but we cannot assume that the interaction terms do not matter.[12]

Table 4: Regression on the probability of choosing the RPS mechanism.

	Full Sample		Knowledge		Ignorance	
			Mechanism			
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	0.564	0.539	0.581	0.626	0.545	0.538
	(0.046)	(0.081)	(0.063)	(0.074)	(0.068)	(0.089)
$RPS \succeq CT$	0.248****	0.258****	0.240***	0.256***	0.259***	0.260***
	(0.064)	(0.063)	(0.089)	(0.087)	(0.092)	(0.093)
Knowledge		0.086				
		(0.095)				
Ability		0.012		-0.103		0.013
		(0.096)		(0.090)		(0.096)
Knowledge & Ability		-0.116				
		(0.132)				
Estimator	OLS	OLS	OLS	OLS	OLS	OLS
N	197	197	101	101	96	96
Adjusted R^2	0.062	0.055	0.053	0.055	0.063	0.053

(* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001)

We observe a similar phenomenon in experiment 2. Figure 1 shows a wide range of incorrect beliefs. In particular, only 4.5% of the subjects report the only correct belief of a 1/2 probability of winning in the MP (the horizontal line at MP Belief = 0). Regarding the choices made, it is clear that the decision of a large number of subjects is not only based on their beliefs on the expected monetary payoff of the two mechanisms. If it were the case, the choices of allocation mechanisms would be cleanly divided between the top-left and bottom-right by the red line.

The beliefs may have some influence on the choices made by subjects, as shown in Table 5. Simply being above or below the red line does not significantly influence the choice made, but being further away from this line may, as shown by the first regression. The explanatory power of these regressions is very small, however. It suggests that beliefs are not the main reason why subjects chose one mechanism over the other in experiment 2.

3.1.2 Discussion

RPS is significantly preferred to MP as an allocation mechanism in experiment 1, but not in experiment 2. We have shown above that some subjects mainly base their decision on their probability of winning, in particular in experiment 1. For other subjects however, it is not

[12]The results including all the interaction terms are not shown, but available by asking the authors.

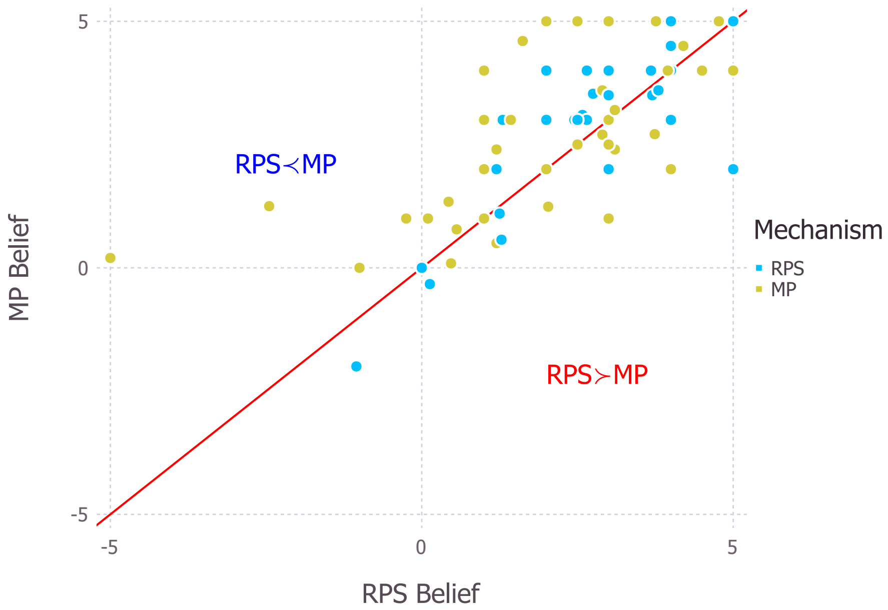


Figure 1: Beliefs and chosen allocation mechanism in Experiment 2.

Table 5: Influence of the beliefs on the choice of the allocation mechanism in experiment 2.

	Mechanism	
	(1)	(2)
(Intercept)	0.450	0.413
	(0.056)	(0.063)
Beliefs $(RPS - MP)$	0.058*	
	(0.035)	
Beliefs $(RPS \succ MP)$		0.049
		(0.117)
Estimator	OLS	OLS
N	89	89
Adjusted R^2	0.010	-0.009

(* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001)

Table 6: Classification of the comments.

	Experiment 1	Experiment 2	p-value[1]
Preference	36%	36%	0.69
Probability	48%	44%	1.0
Sample	153 (out of 197)	72 (out of 89)	

[1] P-value of the Fisher Exact Test.

the case: a majority of subjects knowing that their probability of winning in experiment 1 is lower than 1/2 still choose the RPS. Experiment 2 gives us a tentative explanation of the reasons behind this revealed preference: some people do not like to be the subject of an explicit lottery when an alternative allocation mechanism based on criteria is available. It is possible, however, to transform an explicit lottery by giving subject payoff-irrelevant decisions to make, as in experiment 2. In that case, the aggregate preference for RPS of experiment 1 disappears.

In addition to the strategies and choice of mechanism in experiment 1 and 2, we have asked some demographic questions. These characteristics do not have an impact on the choice of mechanism. In contrast to other situations involving risk and uncertainty (Charness and Gneezy, 2012), gender, in particular, does not play a significant role in our setup (see Appendix G.2).

3.2 Stated Preferences

We have also asked people why they chose each alternative. We have categorized their answers into two groups. The first group is composed of subjects stating that they prefer one mechanism over the other, for various reasons, and that it influenced their choices. The second is composed of subjects stating that their choice was driven by their probability of winning. Not all comments can be neatly divided into these two categories, and some belong to both. Table 6 summarizes how many subjects are classified as having their preference or the probability of winning driving their choice. The comments that could not be cleanly interpreted as belonging to one or the other category are left out of this analysis. The proportion of people claiming to be influenced by probabilities or preference for one mechanism over the other are strikingly similar in the two experiments (and not significantly different). Importantly for our purpose, not all subjects report being influenced by the probability, and a significant proportion of subjects express a taste for one mechanism or another. Among the 24 subjects for which we can tease out what mechanism they prefer in experiment 1, 88% prefer the RPS. Among the 15 subjects for which we can tease out what mechanism they prefer in experiment 2, 53% prefer the RPS. The size of the samples do not allow us to draw firm conclusions, but all the evidence we have gathered point in the same direction.

Table 7 shows that subjects acted on their stated preferences. When subjects stated that their probability of winning was driving their choices, they chose more often the RPS when they believed or knew that it was better, and the difference is significant. When they report a preference that is not based on the probability of winning, they chose RPS much more. It shows that many people intrinsically prefer the RPS to MP as an allocation mechanism in experiment 1. We also see that this difference is not significant at the 10% level for those players who believe/know that their probability of winning RPS is higher than 1/2. In that case, the intrinsic preference for RPS and the higher probability of winning indeed go in the same direction. We see a similar pattern in experiment 2, but at a lower level. We further explore the stated preferences in Appendix E. Subjects who did not state that the probability of winning was the main factor influencing their choice chose the RPS more often than those who said so.

Table 7: Proportion of the RPS choice.

(a) In Experiment 1

$RPS \succeq MP$	No	Yes	p-value
Probability	30%	70%	0.001
Not Probability	70%	90%	0.076
p-value	< 0.001	0.118	

(b) In Experiment 2

$RPS \succeq MP$	No	Yes	p-value
Probability	8%	38%	0.17
Not Probability	63%	62%	1.0
p-value	0.003	0.536	

(P-value of the Fisher exact test)

4 Discussion

The main result of our experiments is that a majority of people (around 55% of our sample) dislike their fate being the subject of an explicit lottery (a coin toss), and prefer to be subject to a set of criteria (how well their choice of RPS strategies fare against a player drawn at random from a database). Our interpretation is that those subjects want to be able to tell themselves they are in control and do not want to attribute their possible success to luck. They are willing to pay a premium for that. A minority of subjects (around 25%) seem to be only influenced by the probability of winning. The rest (around 20%) prefer the explicit lottery over the set of criteria. An interpretation can be that those subjects perceive RPS to be as random as a lottery and prefer the simpler and perhaps more familiar coin toss (Eliaz and Rubinstein, 2014).

This majority preference for criteria disappears if the explicit lottery is re-framed to involve payoff-irrelevant decisions. The willingness to interpret a set of criteria as being different from a lottery is strong. It seems to extend to something as obviously random as the choice of strategies in a repeated matching pennies game against a computer drawing head and tail with equal probability. With the RPS and MP, it is possible to say that we deserve being in a good position because we made the right choices. Explicit lotteries are harder to rationalize than such “non-random” mechanisms.

There may be other motives than not liking randomness to choose RPS over CT in the first experiment. First, subjects may want to know the results of their choices, to discover how well they played in the RPS. However, the fact that the preference for RPS disappears in experiment 2 is difficult to reconcile with this explanation, unless we take the view that people are equally interested to know how well they “played” at guessing how head and tails are drawn by a random device. It is, however, a point we would like to explore in the future, for instance, by varying the level of effort players have to put into the criteria used by our mechanism.

Second, subjects may believe that it is harder for us to manipulate the result of their choices in the RPS compared to the CT. Again, this explanation holds only if one argues that it is also easier for us to manipulate a coin toss drawn by the computer (experiment 1) than five coin tosses (experiment 2). The fear of manipulation is a valid criticism of our experiment: it largely relates to our subjects trusting our algorithms, or at least not trusting one mechanism more than the other. This weakness also extends to the kind of real-world mechanisms we aim to inform with our experiment, typically designed and operated by government agencies. For instance, according to a recent Gallup World Poll,[13] 31% of the UK citizens, 41% of the French citizens and 59% of the US ones answer “no” to whether they believe elections in their country are honest. We believe that the question of how transparent mechanisms can be made is thus an important one, that we wish to explore by making the randomness based on the result of something we visibly cannot control.

A final comment relates to the specificity of the subjects we drew from the AMT sample. It is well-known that AMT subjects behave in a way that is consistent with a standard laboratory

[13] “Faith in Elections in Relatively Short Supply in U.S.,” RJ Reinhart, Gallup, February 13, 2020.

setting (Horton *et al.*, 2011), and allow studying behaviour in an arguably more externally valid way as the panel is more diverse (see for instance Rand *et al.*, 2012; Kuziemko *et al.*, 2015). We would, however, be interested to know if, for instance, the share of people displaying a preference for criteria remains similar in a different setting.

References

Abdulkadiroglu, A., Angrist, J. D., Narita, Y. and Pathak, P. A. (2019). Breaking ties: Regression discontinuity design meets market design.

Agranov, M. and Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, **125** (1), 40–68.

Aldashev, G., Kirchsteiger, G. and Sebald, A. (2017). Assignment procedure biases in randomised policy experiments. *The Economic Journal*, **127** (602), 873–895.

Anderson, S. and Baland, J.-M. (2002). The economics of roscas and intrahousehold resource allocation. *The Quarterly Journal of Economics*, **117** (3), 963–995.

Arnosti, N. and Shi, P. (2020). Design of lotteries and wait-lists for affordable housing allocation. *Management Science*.

Avin, S. (2019). Mavericks and lotteries. *Studies in History and Philosophy of Science Part A*, **76**, 13–23.

Barnett, A. G. (2016). Funding by lottery: political problems and research opportunities. *MBio*, **7** (4), e01369–16.

Bartling, B., Fehr, E. and Herz, H. (2014). The intrinsic value of decision rights. *Econometrica*, **82** (6), 2005–2039.

Besley, T., Coate, S. and Loury, G. (1993). The economics of rotating savings and credit associations. *The American Economic Review*, pp. 792–810.

Bobadilla-Suarez, S., Sunstein, C. R. and Sharot, T. (2017). The intrinsic value of choice: The propensity to under-delegate in the face of potential gains and losses. *Journal of risk and uncertainty*, **54** (3), 187–202.

Bouacida, E. (2019). Eliciting choice correspondences a general method and an experimental implementation.

Brezis, E. S. (2007). Focal randomisation: An optimal mechanism for the evaluation of r&d projects. *Science and Public Policy*, **34** (10), 691–698.

Brownback, A. and Kuhn, M. A. (2019). Understanding outcome bias. *Games and Economic Behavior*, **117**, 342–360.

Cettolin, E. and Riedl, A. (2019). Revealed preferences under uncertainty: Incomplete preferences and preferences for randomization. *Journal of Economic Theory*, **181**, 547–585.

Charness, G. and Gneezy, U. (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization*, **83** (1), 50–58.

Chen, D. L., Schonger, M. and Wickens, C. (2016). otreean open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, **9**, 88 – 97.

Cole, S., Simon, G. A. *et al.* (1981). Chance and consensus in peer review. *Science*, **214** (4523), 881–886.

Deaton, A. and Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, **210**, 2–21.

Donovan, J. L., de Salis, I., Toerien, M., Paramasivan, S., Hamdy, F. C. and Blazeby, J. M. (2014). The intellectual challenges and emotional consequences of equipoise contributed to the fragility of recruitment in six randomized controlled trials. *Journal of clinical epidemiology*, **67** (8), 912–920.

Dwenger, N., Kübler, D. and Weizsäcker, G. (2018). Flipping a coin: Evidence from university applications. *Journal of Public Economics*, **167**, 240–250.

Eliaz, K. and Rubinstein, A. (2014). On the fairness of random procedures. *Economics Letters*, **123** (2), 168–170.

Erdil, A. and Ergin, H. (2008). What's the matter with tie-breaking? improving efficiency in school choice. *American Economic Review*, **98** (3), 669–89.

Fang, F. C. and Casadevall, A. (2016). Research funding: The case for a modified lottery. *mBio*, **7** (2), e00422–16.

Fienberg, S. E. (1971). Randomization and social affairs: the 1970 draft lottery. *Science*, **171** (3968), 255–261.

Fox, C. R. and Tversky, A. (1995). Ambiguity aversion and comparative ignorance. **110** (3), 585–603, publisher: Oxford Academic.

Frank, R. H. (2016). *Success and luck: Good fortune and the myth of meritocracy*. Princeton University Press.

Graves, N., Barnett, A. G. and Clarke, P. (2011). Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. *Bmj*, **343**, d4797.

Greenberg, D. S. (1998). Chance and grants. *The Lancet*, **351** (9103), 686.

Heath, C. and Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. **4** (1), 5–28.

Heckman, J. J. (2020). *Randomization and Social Policy Evaluation Revisited*. Tech. rep.

Horton, J. J., Rand, D. G. and Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental economics*, **14** (3), 399–425.

Keren, G. and Teigen, K. H. (2010). Decisions by coin toss: Inappropriate but fair. *Judgment and Decision Making*, **5** (2), 83.

Kesten, O. and Ünver, M. U. (2015). A theory of school-choice lotteries. *Theoretical Economics*, **10** (2), 543–595.

Kuziemko, I., Norton, M. I., Saez, E. and Stantcheva, S. (2015). How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, **105** (4), 1478–1508.

Leib, E. J. and Galoob, S. (2016). Fiduciary political theory: A critique. *Yale Law Journal*, **125** (7).

Lilford, R. J. and Jackson, J. (1995). Equipoise and the ethics of randomization. *Journal of the Royal Society of Medicine*, **88** (10), 552.

Loewenstein, G. and Issacharoff, S. (1994). Source dependence in the valuation of objects. *Journal of Behavioral Decision Making*, **7** (3), 157–168.

Oberholzer-Gee, F., Bohnet, I. and Frey, B. S. (1997). Fairness and competence in democratic decisions. *Public Choice*, **91** (1), 89–105.

Pathak, P. A., Rees-Jones, A. and Sönmez, T. (2020). *Immigration Lottery Design: Engineered and Coincidental Consequences of H-1B Reforms*. Tech. rep., National Bureau of Economic Research.

Petticrew, M., McKee, M., Lock, K., Green, J. and Phillips, G. (2013). In search of social equipoise. *bmj*, **347**, f4016.

Pier, E. L., Brauer, M., Filut, A., Kaatz, A., Raclaw, J., Nathan, M. J., Ford, C. E. and Carnes, M. (2018). Low agreement among reviewers evaluating the same nih grant applications. *Proceedings of the National Academy of Sciences*, **115** (12), 2952–2957.

Qiu, J. and Ong, Q. (2017). Indifference or indecisiveness: a strict discrimination.

Rand, D. G., Greene, J. D. and Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, **489** (7416), 427–430.

Roumbanis, L. (2019). Peer review or lottery? a critical analysis of two different forms of decision-making mechanisms for allocation of research grants. *Science, Technology, & Human Values*, **44** (6), 994–1019.

Schlag, K. and Tremewan, J. (2020). Simple belief elicitation: an experimental evaluation.

Singer, P., Baker, A. and Haushofer, J. (2019). Are randomized poverty alleviation experiments ethicals? *Project Syndicate*.

Toroyan, T., Roberts, I. and Oakley, A. (2000). Randomisation and resource allocation: a missed opportunity for evaluating health care and social interventions. *Journal of Medical Ethics*, **26** (5), 319–322.

Vermeule, A. (2015). Rationally arbitrary decisions in administrative law. *The Journal of Legal Studies*, **44** (S2), S475–S507.

APPENDIX

A Lotteries in Allocation Problems

In this section, we provide a cursory review of the use of lotteries in real-world allocation problems and the debates around it. In the first part, we discuss the idea of *equipoise*, the basic ethical principle behind the use of randomization in medicine. We then show how a similar idea is used to justify lotteries outside of the medical field. In the second part, we discuss the difference between using lotteries to make a decision and being the subject of a lottery.

A.1 Equipoise and the Need Exhaust All Possible Criteria

In the medical literature, randomization is seen as ethical if a state of *equipoise* is reached: a consensus that none of the possible outcomes of the lottery is *ex ante* better than the others (Lilford and Jackson, 1995). For this reason, the medical profession often reports a difficulty to recruit subjects for Randomized Control Trials (RCTs). One explanation is that it is very difficult for a doctor not to seek an additional rationalization instead of declaring two possible treatments to be *ex ante* equivalent (Donovan *et al.*, 2014). In our context, the condition of equipoise corresponds to the idea that there is no allocation problem, as all the options are equivalent *ex ante*. Equipoise is, however, not always necessary. When resources are scarce – a new medicine available in limited quantity, HIV in the developing world has been a prominent example – randomization is an acceptable allocation mechanism. The difficulty is to make people acknowledge that an allocation mechanism is necessary and that no criterion is fairer than the use of a random device (Lilford and Jackson, 1995; Toroyan *et al.*, 2000).

RCTs are also common in economics. Here, equipoise would be genuine uncertainty about the relative benefits of different interventions (Petticrew *et al.*, 2013). Most RCTs in economics, however, use a random device to provide a treatment with expected positive effects to a subgroup of the population. The rest of the population should not benefit, and even sometimes suffer from the absence of treatments (Aldashev *et al.*, 2017; Deaton and Cartwright, 2018; Heckman, 2020). For this reason, RCTs are often criticized as being unethical. As in the medical literature, proponents of RCTs defend the use of a random device as the fairest way to allocate an intervention when resources are scarce (Singer *et al.*, 2019).

What makes lotteries socially acceptable is thus the conviction that there is no better way to decide. Public opposition to the use of random devices is often grounded on the case that other criteria would be fairer, or even that no allocation mechanism is necessary.[14] Similarly, most lotteries used to solve real-world allocation problems are accepted because, first, the allocation problem exists and that, second, there is no better rationalization for the decision. In the case of the military draft for instance: "When a number of people fulfil all the requirements of, or qualifications for, a particular position or duty, the notion of a lottery or random selection allows each individual the same chance of attaining that position or being selected for that duty (Fienberg, 1971)."[15] Another prominent real-world use of lotteries for allocation purposes is the Rotating Savings and Credit Associations (Besley *et al.*, 1993; Anderson and Baland, 2002). ROSCAs are an informal financial institution in which a joint saving is regularly allocated at

[14] Justifying his opposition to lotteries in school allocation, future UK Secretary of State for Education Michael Gove claims in a March 2008 BBC interview to be "Completely opposed to lotteries because lotteries aren't a way of solving this problem. With a lottery all you're basically trying to do is come up with an ingenious scheme to allocate a small number of good school places (...) [We] need to ensure that (...) parents get the sort of school that they want for the children." To the best of our knowledge, there is no high profile case against the use of a random allocation procedure explicitly stating that the parties would have higher chances with an objective set of criteria.

[15] Similarly, Lyndon Johnson justified the Vietnam draft as an "equal and uniform treatment for all men *in like circumstances*" (our emphasis, Congressional Record of the US senate. March 6, 1967, p.5479.)

random to one of the members. They are particularly prevalent in the developing world, precisely because its members acknowledge the scarcity of resources and the need for an allocation mechanism.

A.2 Using Lotteries to Decide VS Being the Subject of a Lottery

As mentioned in the introduction, our result of an aversion to being the subject of a lottery contrasts with the well-understood idea of "preference for randomization": When indifferent, individuals often choose to use a lottery to make their choice. These two very different effects suggest that what makes a lottery socially acceptable can also be the extent to which those choosing the allocation mechanism are at the same time the subjects of the mechanism.

Perhaps one of the most famous lotteries in the world is the "Green card lottery." Since 1992, the US Diversity Immigrant Visa program awards permanent residency to around 50,000 people a year, combined with criteria based on the origin of the applicant and mild education or employment requirements. It is, however, a widely criticized program, and its opponents – among which the current US president – argue that it is unfair and there should be additional criteria to distinguish between applicants.[16] For the short-term skilled migration programs H-1B, a random tie-breaking rule is used to choose applicants among those who qualify, but the lottery part of the program is not widely advertised (Pathak *et al.*, 2020). A common characteristic of those two lotteries is that the participants are, by definition, not US citizens. How they perceive their relative ability to perform under different procedures has, therefore, no impact on the social acceptability of the random device by the US public.

Lotteries are also used in the allocation of affordable housing in several US cities. The use of the randomization device is typically combined with an assessment of needs so that all the perceived "fair" criteria have been exhausted. The simplest and most efficient randomization device, a common lottery to determine priority for all items (Arnosti and Shi, 2020) is, however, much less used than more complex individual lotteries combined with additional criteria. This case is, again, one in which taxpayers decide the best way to allocate money among a small part of the population. It is not a decision taken directly by those who could potentially benefit from it, and how acceptable the lottery is for taxpayer matters at least as much as how acceptable it is for the beneficiaries.

B Framing Treatments

This section summarizes the difference between the "ability" and the "chance" treatments. In the description of the RPS game (stage 2 of the timing), we used the following formulations:

1. One as a child's play of chance (Chance)

   > Rock-paper-scissors (RPS) is an old child play originating from ancient China. [...]

2. One as a game of ability and strategic thinking (Ability)

   > Rock-paper-scissors (RPS) is an old game of strategy originating from ancient China. [...] RPS is also a well-studied game in biology, psychology, and artificial intelligence. In international competitions, some players consistently outperform the others. Computer scientists have produced algorithms able to exploit the predictable behaviour of human players and win more often against them.

[16] The US President Donald Trump declared in 2018 that he would "only sign an immigration deal that cancels the diversity visa lottery program" as "under the diversity visa program, countries give us their worst people, they put them in a bin" (all quotes from Rafael Bernal, "What is the diversity visa Lottery?", The Hill, 01/28/18).

Then, at the time of making the actual decision (stage 3 of the timing) we added the following sentence in the Ability treatment:

> Your probability of winning the RPS depends on how well your chosen actions performed against past RPS players.

C Description of the RPS Database

Table 8 summarizes the strategies of the 2,500 players in the Roshambo sample and the subjects in our experiments. The latter perform "worse" than the players in Roshambo, as shown by their quite unbalanced choice between the three different actions, even more heavily biased towards the choice of "rock." It is also illustrated by the fact that the best pure strategy has very high odds of winning in the two experiments compared to the Roshambo data set. This result is reassuring in the sense that the bias in favour of "rock" is easy to find when looking for the best strategies to win RPS on any Internet search engine. It seems this is not something our subjects tried to do. We denote by Worst (resp. Best) profile the sequence of strategies that give the lowest (resp. highest) probability of winning against a sample.

Table 8: Descriptive statistics of the strategies played in the different RPS games.

	Roshambo	Experiment 1	Experiment 2
Rock	34.56%	42.03%	44.49%
Paper	33.95%	36.65%	33.26%
Scissors	31.49%	21.32%	22.25%
Best Profile	(P, S, P, P, P)	(P, P, P, P, P)	(P, P, S, P, P)
Probability Best	55.00%	79.19%	78.65%
Worst Profile	(R, R, R, S, S)	(R, R, S, R, S)	(R, R, R, R, S)
Probability Worst	44.80%	24.37%	24.16%

D The 50% Threshold

Figure 2 plots the real probability of winning against the elicited beliefs in experiment 1 in the ignorance treatment, as well as the allocation mechanism chosen. The figure shows that subjects are not very good at predicting their probability of winning, otherwise, they would all be on the red line. 42.7% of subjects believe that their probability of winning the RPS is higher than 50%, but 66% of them choose the RPS. The beliefs are still correlated with choices. There are proportionally more blue dots on the right than on the left of the blue line (80.5% versus 54.5%).

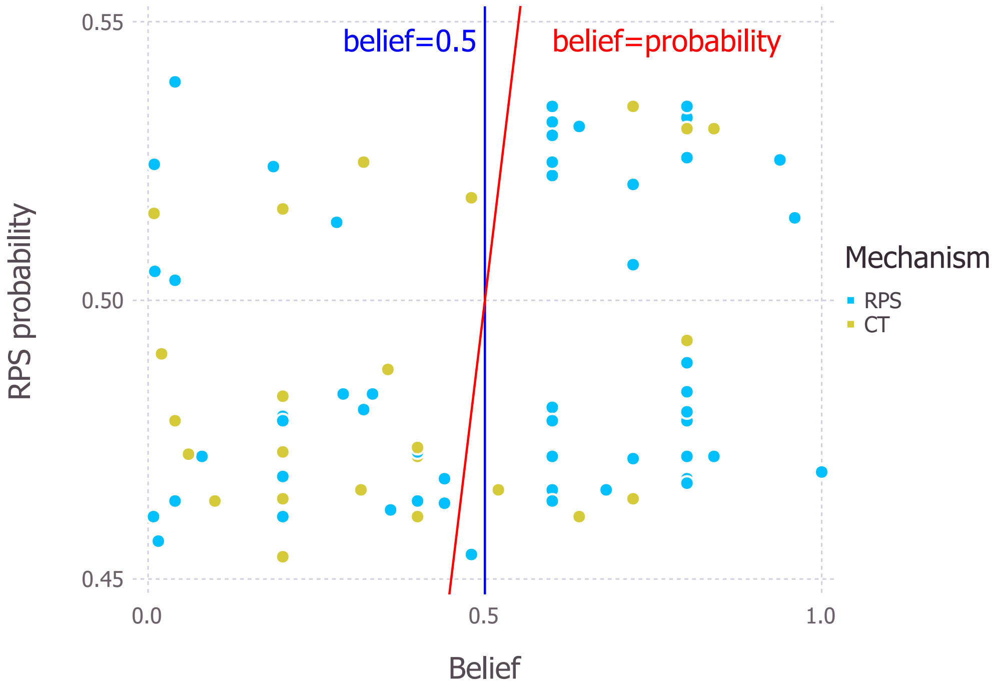


Figure 2: Beliefs and decisions when the probability is unknown.

E Additional Regressions in Experiment 1

E.1 Interaction Terms

Table 9 shows the regression of Table 4 using all interactions terms. None of the additional interaction terms is significant, which justifies the use of a shorter model in Section 3.1.1.

E.2 Using Distance in Beliefs

Table 10 is the exact counterpart of Table 4 but using the distance instead of a dummy for the preference. We find that the results are robust to this modification. The magnitude of the coefficient differs between the knowledge and ignorance treatment because the real difference in probability is an order of magnitude smaller than the believed one (0.023 vs 0.28 in absolute value). In short, the effect of beliefs is robust to how we specify them.

E.3 Adding Stated Preferences

Table 11 shows that stated preferences for the probability do matter. Subjects who care about the probability choose less often the RPS, compared to the baseline. As we said earlier, knowing the probability and having a probability of winning the RPS that is higher than 50% also drive subjects to choose the RPS. This effect is mainly significant for subjects who mention that probability drives their choices.

F Ambiguity Aversion

Table 12 presents the probability of choosing RPS in the first experiment depending on whether or not subjects display ambiguity aversion.

Table 9: Regression capturing the effect of believing or knowing that the probability of winning the RPS is higher than the probability of winning the Coin Toss in experiment 1.

	Full Sample	Knowledge	Ignorance
		Mechanism	
(Intercept)	0.545	0.600	0.545
	(0.108)	(0.084)	(0.108)
$RPS \succeq CT$	0.244*	0.338***	0.244*
	(0.144)	(0.105)	(0.145)
Knowledge	0.055		
	(0.137)		
Ability	-0.000	-0.044	-0.000
	(0.140)	(0.129)	(0.140)
$RPS \succeq CT$ & Knowledge	0.093		
	(0.178)		
$RPS \succeq CT$ & Ability	0.029	-0.154	0.029
	(0.189)	(0.171)	(0.189)
Knowledge & Ability	-0.044		
	(0.190)		
$RPS \succeq CT$ & Knowledge & Ability	-0.183		
	(0.255)		
Estimator	OLS	OLS	OLS
N	197	101	96
Adjusted R^2	0.043	0.052	0.043

(* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001)

Table 10: Regression capturing the effect of the distance between the (believed or real) probability of winning the RPS and the Coin Toss in experiment 1, using a restricted regression model.

	Full Sample		Knowledge		Ignorance	
			Mechanism			
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	0.685	0.683	0.714	0.770	0.692	0.681
	(0.033)	(0.073)	(0.044)	(0.058)	(0.047)	(0.073)
$RPS - CT$	0.391***	0.405***	4.922***	5.263***	0.369**	0.372**
	(0.148)	(0.148)	(1.706)	(1.669)	(0.148)	(0.149)
Knowledge		0.028				
		(0.097)				
Ability		0.021		-0.108		0.019
		(0.097)		(0.090)		(0.097)
Knowledge & Ability		-0.090				
		(0.135)				
Estimator	OLS	OLS	OLS	OLS	OLS	OLS
N	197	197	101	101	96	96
Adjusted R^2	0.029	0.017	0.059	0.063	0.047	0.037

(* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001)

Table 11: Full model regression taking into account stated preferences. Probability is a dummy with value 1 when the subjects said that probability matter for their choice.

	Mechanism
	(1)
(Intercept)	0.692****
	(0.085)
Probability	-0.359***
	(0.127)
$RPS \succeq CT$	0.208
	(0.128)
Knowledge	0.016
	(0.122)
Probability & $RPS \succeq CT$	-0.255
	(0.228)
Probability & Knowledge	-0.088
	(0.179)
$RPS \succeq CT$ & Knowledge	-0.027
	(0.212)
Probability & $RPS \succeq CT$ & Knowledge	0.640**
	(0.311)
Estimator	OLS
N	153
Adjusted R^2	0.213

Note: Using the distance $RPS - CT$ yield a similar table, but has been omitted for brevity.
(* p<0.1, ** p<0.05, *** p<0.01, **** p<0.001)

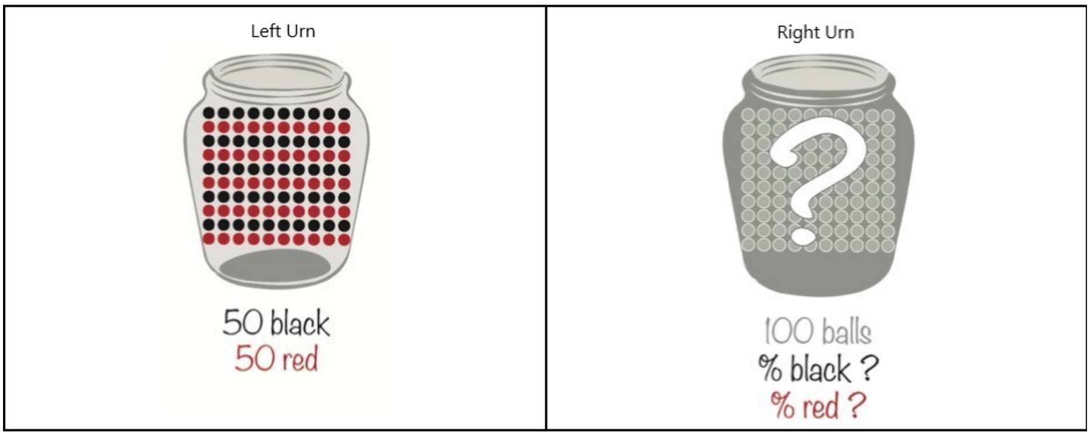
Table 12: Ambiguity Aversion on the Mechanism Chosen

	Ambiguity Averse	Not AA	p-value[1]
Knowledge	61%	73%	0.29
Ignorance	57%	72%	0.18
Everyone	59%	72%	0.07

[1] P-value of the Fisher exact test.

As we would expect, ambiguity averse subjects choose more often the coin toss, which has a known and immediate probability of winning.

To elicit ambiguity aversion, subjects made the following incentivized choice:



> The two urns in the image above contain red and black balls. We will draw one ball from one of these urns. The left urn contains 50 red and 50 black balls. The right urn contains 100 red and black balls, but in a proportion entirely unknown to you. You can choose the urn from which the ball is drawn and the colour to bet on. You will earn \$0.10 if you bet on the colour drawn (and \$0.00 otherwise).
>
> What urn do you choose?
>
> What ball colour do you choose?

We classified as ambiguity averse the subjects choosing the left urn. It means that we may be classifying as ambiguity averse subjects who are ambiguity neutral. So our elicitation of ambiguity aversion is an upper bound of ambiguity averse subjects. The effect of ambiguity aversion is therefore limited at best, as our little significant p-value shows.

G Demographics

G.1 Descriptive Statistics

We have asked subjects for 4 main demographic questions: gender, country of residence, employment status and age. The questions are self-reported and not incentivized. The majority of our subjects comes from the United States (68.0% in experiment 1 and 73.0% in experiment 2). The second highest country of residence is India, which represents 24.9% and 15.7% of the sample. The majority are employed (79.7% in experiment 1 and 77.5% in experiment 2). The majority are between 25 and 40 years old (75.6% and 76.4%). The proportion of employed residents of the USA between 25 and 40 years old represents 39.6% and 44.9% of the sample, respectively. If we gather Indian and USA residents, the proportions become 58.8% and 55.1%. The two samples are similar from a demographic standpoint. The difference in the choice of the allocation mechanism is not because of a difference in the sample. The relative homogeneity of the sample does not allow us to investigate the effect of demographic characteristics on choices.

The only demographic characteristic which is sufficiently heterogeneous to allow for an investigation is gender, and we report it in the next Section.

G.2 Gender

Table 13 presents the probability of choosing RPS in the two experiments depending on the self-reported gender of the participants. Only one subject picked the option not to do so, in the

Table 13: Gender on the Mechanism Chosen

	Male	Female	p-value[1]
Experiment 1	69%	61%	0.37
N	135	61	
Experiment 2	47%	38%	0.54
N	47	42	

[1] p-value of the Fisher exact test.

first experiment. The subject is excluded from this analysis, There is no significant difference between genders. If we add gender to the regression of Tables 4 or Table 5, the effect is not significant.